ANISOTROPY FACTOR ESTIMATION FROM BACKSCATTERED Q ELEMENTS OF STOKES VECTORS.

Julie Falconet[1], Raphaël Sablong[1], Emmanuel Perrin[1] and Hervé Saint-Jalmes[2]

[1]CREATIS – LRMN CNRS UMR 5220 INSERM U630 Université de Lyon, INSA de Lyon, Université Claude Bernard Lyon 1, 69616 Villeurbanne Cedex, France
[2]PRISM Villejean, Université Rennes 1, Faculté de Médecine, CS 34317, 35043 Rennes Cedex, France, Département d'Imagerie, Centre Eugène Marquis, CS44229, 35042 Rennes Cedex, France

ABSTRACT

Optical characterization of biological tissues is of real interest to improve medical diagnosis and in particular in the detection of precancerous tissues. We propose a new non-invasive method allowing the estimation of the anisotropy factor. This method is based on the image analysis of the Q-element of Stokes vector backscattered from the turbid medium. These Q-element images show specific patterns depending on *g*. Therefore the use of Fourier Descriptors (*FD*) on simulated data, to discriminate the specific geometrical features of the Q-element, enabled us to determine a linear relation between the anisotropy factor and six *FD*. This method was applied on experimental data obtained with calibrated solutions. The anisotropy factor was estimated with a maximum relative error of 13 %.

Index Terms— light scattering, linearly polarized light, scattering media, anisotropy factor, Fourier Descriptors.

1. INTRODUCTION

Optical imaging provides a non invasive way to analyze turbid media, which is of real interest for investigating biological tissues for diagnosis purposes [1]. These media can be characterized by their absorption coefficient μ_a, reduced scattering coefficient μ'_s, and anisotropy factor *g*. For biological media, light scattering is due to the heterogeneity of cellular and extra-cellular structures constituting tissues [2]. The characteristics of these structures are potentially related to the pathological state of the medium, *e.g.* cancer may cause morphological changes [3]. As the size of scattering elements directly influences the anisotropy factor, the estimation of *g* is expected to improve the tissue characterization. However measurements of the scattering coefficient μ_s, or the anisotropy factor *g*, did not give satisfying results so far [4, 5].

The illumination of a medium with linearly polarized light provides specific patterns in the backscattered images, strongly depending on *g* [6]. In this paper, we propose a method to find a linear model relating features of the Q-element of Stokes vector images and the anisotropy factor.

As a first step a feasibility study was carried out [7]. Correlation coefficients calculated on Monte Carlo simulations of different media showed that *g* can be deduced from the images. Shape descriptors (called Fourier Descriptors [8]) are used to increase the quality of the results. But these Fourier Descriptors are too numerous to obtain a simple relation with *g*, and the relation between *g* and each descriptor is not bijective.

Thus in this paper, as a second step, a statistical study based on an analysis of variance is processed in order to select a reduced set of significant Fourier Descriptors. A multiple regression is then performed on this set of parameters to find a linear relation between some Fourier Descriptors and the anisotropy factor.

This relation is used to estimate the *g* values of calibrated solutions for which experimental images are obtained. Mie theory enables to calculate the expected g values. Thus this latter can be compared to the experimental value. The anisotropy factors of the calibrated media are chosen in the range $\{0.6; 1\}$, which corresponds to the case of biological tissues [2]. Finally results are described and discussed.

2. IMAGE ANALYSIS

The quantitative analysis of the different backscattered patterns is expected to point out a reduced set of "relevant features" that are significantly sensitive to *g*. Thus the *FD* analysis of circular profiles seems to be an interesting approach for such a purpose [8]. A closed contour *c(n)* can be considered as a 2π-periodic function. It can thus be decomposed in its Fourier series, which coefficients noted *FD(k)* are called Fourier Descriptors [9].

The *FD* are calculated at different radii, thus they are denoted as:

$$FD_r(k) = \frac{1}{N_r}\sum_{n=0}^{N_r-1} c_r(n)\exp\left(-2\pi jk\frac{n}{N_r}\right), \qquad (1)$$

N_r being the number of points of $c_r(n)$, at the radius *r*.

The *FD* represent the frequential components of the contour: low indices coefficients represent low frequencies or general shape of the outline, and higher indices coefficients represent high frequencies or details of the shape. So, the *FD* will describe the relevant features of the different lobes that appear in the images of backscattered light.

Fig. 1 shows different simulated images. The corresponding *FD* moduli are represented as pixels intensities, with *FD* indices on the columns, and different radii *r* on the lines of the figures. An obvious link between the number of lobes in the images and the index of the most significant *FD* can be observed. More precisely, the maximum number of observed lobes being 4 for $g = 0.006$, the maximum index of *FD* that seems to be significant is 5, whereas 8 observed lobes for $g = 0.9$ (Fig. 1), give a maximum index of *FD* equal to 9. In Fig. 1, only the low indices of *FD* are shown, and only positive indices, as the coefficients are symmetric.

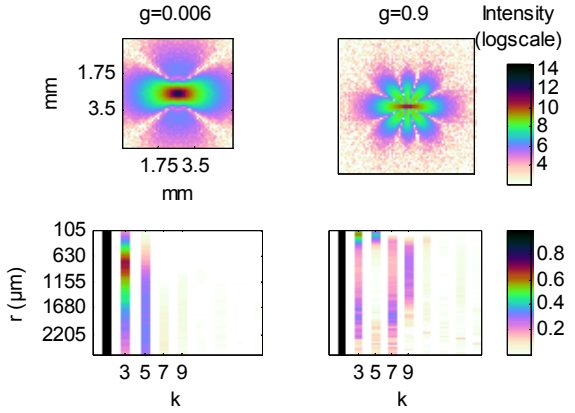

Fig. 1. Simulated images (logarithm of the absolute value) of linearly polarized light backscattered by media of $g = 0.006$ and $g = 0.9$ and the modulus (coded with the scale indicated on the right of the last image) of the corresponding *FD* for different radii. $\mu'_s = 20$ cm^{-1}, $\mu_a = 0.01$ cm^{-1}.

A statistical approach is proposed to find a linear relation between only the most significant *FD* and *g*, in the range of biological media (*g* between 0.6 and 1). The method for obtaining a reduced set of *FD* is presented in two steps: firstly an analysis of variance gives the most significant *FD* (indices and radii), and then a finer selection with a stepwise regression is processed to find the relation between the most relevant *FD* and *g*.

3. SELECTION OF A RELEVANT SET OF FEATURES ON SIMULATED MEDIA

The *FD* set being constituted of N_r different indices over 72 radii, the aim of this step is to find a reduced set of relevant *FD* for determining *g* from backscattered images. As statistical evidence is based on numerous measurements, we choose to use simulated media to get this set of data. In this part, the 2-step methodology of *FD* selection, based on simulated media, is explained.

3.1. Preselection of relevant *FD*

Our aim is to find a criterion sensitive to a variation of *g*, in the range 0.6 to 1, and specific for the determination of a *g* value. The idea is thus to find the *FD* for which the variance is minimal into the group (specificity) and maximal between the different groups (sensitivity).

As a first step, an analysis of variance determines the level of significance of each *FD* (characterized by index and radius). The simulated set of data is constituted of 3 media of 3 different anisotropy factors, with 10 simulations per medium. The anisotropy factors 0.7, 0.8, and 0.9 are typical values of biological media. The reduced scattering coefficient and the absorption coefficient were fixed: $\mu'_s = 20$ cm^{-1} and $\mu_a = 0.01$ cm^{-1}.

The *FD* selected by this analysis of variance have indices 3, 5, 7 and 9. Now watching to the polarization figures it can be observed that 8 lobes appear at most. So the maximum significant index is 9 (the index k corresponds to k-1 distortions [8]). Furthermore the figures have at least 2 symmetry axes [10], implying an even number of lobes, so only odd indices of *FD* are significant. So the results of the analysis of variance are in good agreement with the analysis of the images, showing the relevancy of our approach.

But the threshold given by the analysis of variance is too low: too many *FD* are still selected to obtain a simple relation. A stepwise regression will enable to select the most significant *FD* from the 50 selected by the analysis of variance, and find a linear relation between them and *g*. This second step is explained in the following section.

3.2. Multiple linear regression

In this second step, the *FD* of indices and radii selected in the first step, are used to process a stepwise regression, adding progressively the variables in the model, until optimizing the adjusted R^2. At each step, the most statistically significant variable is added (least p-value), or, if any, the least significant term is removed (p-value became higher than the maximum p-value accepted). The coefficient of each predictor is calculated with the least squares method. Using this method, 6 *FD* were found to adjust R^2 to 0.995. Equation (3) shows the relation between $\hat{g}$ (estimated *g*) and *FD(index, r (in µm))*.

$$\begin{aligned}\hat{g} = {} & 0.874 - 0.179\times FD(3,490) - 0.115\times FD(3,595) \\ & -0.183\times FD(3,665) + 0.312\times FD(9,315) \\ & +0.118\times FD(9,560) + 0.159\times FD(9,770)\end{aligned} \qquad (2)$$

Note that the radii indicated are multiple of 35 which is the resolution of the images. A study showed that changing these radii or decreasing the resolution by calculating $\hat{g}$ for

FD averaged over 3 consecutive radii, do not change the results significantly (results not shown here).

To assess the accuracy of the model, $\hat{g}$ are compared to g, input values of the simulations. Relative errors are calculated for each simulated media. The mean error is 1.3 % and the maximum error is 5.9 % (Fig. 2).

The method allows finding a linear relation between the *FD* and the anisotropy factor of a simulated medium, for a given μ'_s, and give satisfying results on simulations.

In the next section, the relation is applied to images obtained experimentally with calibrated media.

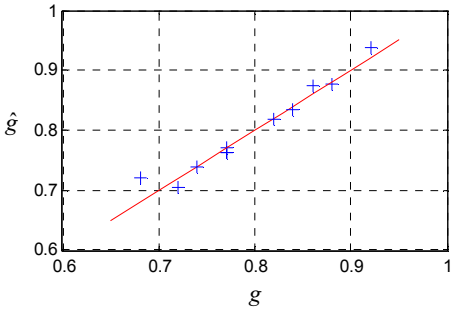

Fig. 2. Estimated anisotropy factors $\hat{g}$ versus theoretical anisotropy factors g. Continuous line is ideal case: $\hat{g} = g$.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section the experimental setup is described and experimental results on calibrated solutions are next presented. The relation obtained in the previous section using simulation is applied to experimental images, and results are discussed.

4.1. Experimental setup

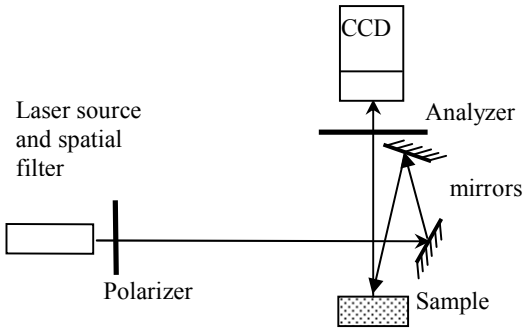

Fig. 3. Scheme of the experimental setup.

A laser diode beam illuminates the medium of interest. A spatial filter is used to provide a Gaussian spot. The resulting spot size is approximately 200 µm and its power is typically some µW. The outgoing beam is horizontal. A couple of mirrors enables to get a quasi vertical beam illuminating the plane surface of the medium. The sample scatters the incoming light. A 2D CCD camera (HX516 *StarLight Xpress Ltd., England*) placed right over the sample captures the light backscattered by the medium. The incident angle of the beam, around 4°, is estimated from the position of the mirrors. This angle avoids detecting specular reflections on the camera. A polarizer is used to increase the polarization rate of the source and an analyzer is placed in front of the camera objective.

4.2. Description of phantoms and experimental images

Suspensions of polystyrene microspheres (Merck®) whose diameter and concentration are precisely indicated by the provider are used as calibrated scattering media. The optical parameters μ'_s and g are calculated by the Mie theory. Seven different suspensions are used, which anisotropy factors in the range of biological media are indicated in Tab. 1. The reduced scattering coefficient is approximately 20 cm^{-1} and the absorption coefficient is closely null.

Images of reflectance of the solutions with parallel polarizers and then crossed polarizers are taken, then the Q-elements of the Stokes' vector is obtained by subtracting the images. Ten images are taken for each sample. Fig. 4 shows Q-elements obtained for some of the solutions.

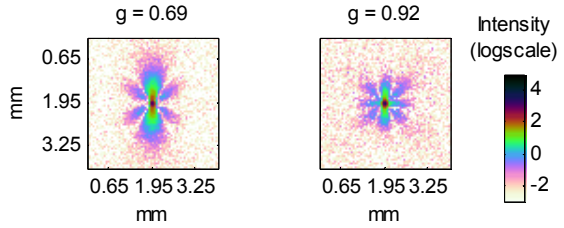

Fig. 4. Experimental images (logarithm of the absolute value) of some calibrated solutions. $\mu'_s = 20$ cm^{-1}, $\mu_a = \mu_{a,water} = 10^{-3}$ cm^{-1}.

4.3. Results

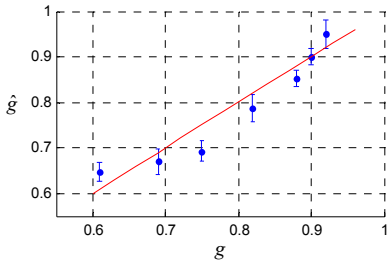

Fig. 5. Estimated anisotropy factors $\hat{g}$ versus theoretical anisotropy factors g (obtained with Mie Theory). The mean relative error is of 4.5 % and the maximum relative error of 13.1 %, with respect to theoretical value. Continuous line is ideal case: $\hat{g} = g$.

The anisotropy factor of each solution is evaluated with the linear relation found in section 3.2. The estimated anisotropy factors are shown in Tab. 1. Fig. 5 plots the mean estimates and the errors as function of the theoretical g values.

4.4. Discussion

The relation obtained with simulated images allows retrieving the anisotropy factor of spheres suspensions with a maximum relative error equal to 13.1 %. A part of this error may be due to the intrinsic *g* variation of the solutions: there is variability in diameter, involving variability in *g*. The standard deviations of theoretical and estimated values are indicated on Tab. 1. The absolute errors of the estimates with respect to theoretical values are also indicated. A correlation can be seen between the standard deviation of the theoretical *g* values and the standard deviation of the estimates. This is shown calculating t-values for two different tests. The first one tests whether the mean of a normally distributed population ($\hat{g}$) has a value specified in a null hypothesis (equal to the theoretical *g*). The results are called t-value 1 in Tab. 1. The second one tests whether the means of two normally distributed populations are equal, i.e. mean theoretical *g* with theoretical standard deviation due to diameter variability, versus mean estimated *g* with standard deviation of estimates. The results are called t-value 2 in Tab. 1.

The first test give t-values higher than 2, the critical value for a 5 % confidence interval, which means $\hat{g}$ is statistically different from theoretical *g*. But considering the second test, taking into account the standard deviation of theoretical *g*, the means are statistically equal. Thus the variability observed for $\hat{g}$ can be explained by the variability of the spheres diameter.

Mean theoretical *g*	Theoretical standard deviation	Mean estimated *g*	Estimates' standard deviation	t-value 1	t-value 2
0.61	0.015	0.646	0.021	5.42	-1.39
0.69	*	0.669	0.029	-2.29	*
0.75	0.005	0.692	0.022	-8.34	2.57
0.82	0.025	0.788	0.029	-3.49	0.84
0.88	0.010	0.852	0.018	-4.92	1.36
0.90	0.017	0.8991	0.018	-0.16	0.04
0.92	*	0.950	0.031	3.06	*

Tab. 1. Variability of theoretical (due to sphere diameter variability) and evaluated *g* values for each solution. t-values 1 and 2 stand for different t-tests. Please refer to text for more comments.
* Variability in diameter not given by the provider.

5. CONCLUSION

In this paper it is first demonstrated that Fourier Descriptors can be useful to increase the ability of backscattered images of turbid media obtained in polarized light to distinguish different anisotropy factors. Then, a method is proposed to find a linear relation between the *FD* and the anisotropy factor. It is first tested on simulations obtained with a Monte Carlo code and then calibrated solutions are used to test the method on experimental data. The anisotropy factors of the tested solutions are retrieved with a maximal relative error equal to 13.1 %.

These results are promising, and should be extended to other media. Indeed the model proposed is validated only for media of same μ'_s, and the μ_a are closely null. Furthermore it is obvious that biological media for example are not monodisperse media as it was the case for the solutions used (there was only a thin dispersion of spheres' diameters). Other teams have worked on the composition of phantoms to be closer to biological tissues [11]. A further study may consist on analyzing bi-disperse media, as a first step to polydisperse media. Finally, here only the Q-element of the outgoing Stokes' vector was considered. But the Mueller matrix of the medium contains other elements which could also be used for the characterization [12].

AKNOWLEDGMENTS

This work has been supported by the Région Rhône-Alpes in the context od project PP3 ["Computer systems dedicated to Medicine and Health"]/I3M ["Multiscale Medical Imaging and modeling: from the small animal to the human being"] of cluster ISLE.

REFERENCES

[1] A. A. Nezhuvingal, L. Yanfang, H. Anumula, and B. D. Cameron, "Mueller matrix optical imaging with application to tissue diagnostics," presented at Laser-Tissue Interaction XIV, San Jose, CA, USA, 2003.

[2] J. R. Mourant, J. P. Freyer, A. H. Hielscher, A. A. Eick, A. Shen, and T. M. Johnson, "Mechanisms of light scattering from biological cells relevant to noninvasive optical-tissue diagnostics," Applied Optics, vol. 37, pp. 3586-93, 1998.

[3] E. Salomatina, B. Jiang, J. Novak, and A. N. Yaroslavsky, "Optical properties of normal and cancerous human skin in the visible and near-infrared spectral range," J. Biomed. Opt., vol. 11, pp. 064026, 2006.

[4] D. J. Faber, F. van der Meer, M. C. Aalders, M. de Bruin, and T. G. van Leeuwen, "Hematocrit-dependance of the scattering coefficient of blood determined by optical coherence tomography," presented at European Congress of Biomedical Optics, München, 2005.

[5] N. Joshi, C. Donner, and H. W. Jensen, "Noninvasive measurement of scattering anisotropy in turbid materials by nonnormal incident illumination," Opt Lett, vol. 31, pp. 936-8, 2006.

[6] A. H. Hielscher, J. R. Mourant, and I. J. Bigio, "Influence of particle size and concentration on the diffuse backscattering of polarized light from tissue phantoms and biological cell suspensions.," Applied Optics, vol. 36, pp. 125-135, 1997.

[7] J. Falconet, R. Sablong, F. Jaillon, E. Perrin, and H. Saint-Jalmes, "Towards optical characterization of biological media: analysis of backscattered images in linearly polarized light, simulations and experiments," presented at Optics and Optoelectronics, Warsaw, Poland, 2005.

[8] C. T. Zahn and R. Z. Roskies, "Fourier Descriptors for Plane Closed Curves," IEEE Transactions on Computers, vol. c-21, pp. 269-281, 1972.

[9] A. Folkers and H. Samet, "Content-based image retrieval using Fourier descriptors on a logo database," presented at 16th International Conference on Pattern Recognition., Quebec City, Que., Canada, 2002.

[10] F. Jaillon and H. Saint-Jalmes, "Description and time reduction of a Monte Carlo code to simulate propagation of polarized light through scattering media," Applied Optics, vol. 42, pp. 3290-3296, 2003.

[11] B. Gélébart, E. Tinet, J. M. Tualle, and S. Avrillier, "Phase function simulation in tissue phantoms: a fractal approach," Pure Applied Optics, vol. 5, pp. 377-388, 1996.

[12] J. S. Baba, C. Jung Rae, A. H. DeLaughter, B. D. Cameron, and G. L. Cote, "Development and calibration of an automated Mueller matrix polarization imaging system," J. Biomed. Opt., vol. 7, pp. 341-9, 2002.